

Application No. 10/033,586

Amendments to the Claims:

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1. (Currently Amended): A machine-implemented method for extrapolating user profile information from user web page access patterns, comprising:

detecting a set of web pages accessed by a test user having an unknown user profile attribute;

mapping at least a subset of said detected web pages to a first data structure, said first data structure representing a web page access pattern of said test user;

comparing said first data structure to a plurality of a second data structure to obtain a comparison result, the plurality of said second data structure representing clusters of web page access patterns of a sample data set of users having a known user profile attribute in common;

evaluating based on said comparison result the plurality of said second data structure and said first data structure to identify a second data structure matching the web page access pattern of the first data structure; and

assigning said unknown user profile attribute of said test user from the matching second data structure to said test user;

wherein the known user profile attribute in common of the sample data [[sets]] set of users corresponds to the unknown user profile attribute of said test user.

2. (Previously Presented): The method of claim 1, wherein said first and second data structures are multi-dimensional vectors, wherein each dimension of said first and said second multi-dimensional vectors corresponds to a different web page, and wherein each dimension of said second data structure corresponds to an average of user multi-dimensional vectors in its corresponding cluster.

Application No. 10/033,586

3. **(Currently Amended):** The method of claim 2, wherein said comparing ~~[[step comprises:]]~~ further comprises determining a distance between said multi-dimensional vectors.

4. **(Previously Presented):** The method of claim 3, wherein said determining further comprises computing a cosine of an angle between said multi-dimensional vectors.

5. **(Previously Presented):** The method of claim 2, wherein said unknown user profile attribute is demographic information.

6. **(Currently Amended):** The method of claim 5, wherein said demographic information is gender information of the test user ~~[[one of gender and age]]~~.

Claims 7-24. **(Canceled)**

Claims 25-27. **(Canceled).**

28. **(Currently Amended):** An apparatus, comprising:

- a memory, said memory adapted to store program code;
- a processor in communication with said memory, said program code capable of programming said processor to perform a method for extrapolating user profile information from user web page access patterns, the method comprising:
 - detecting a set of web pages accessed by a test user having an unknown user profile attribute;

Application No. 10/033,586

mapping at least a subset of said detected web pages to a first data structure, said first data structure representing a web page access pattern of said test user;

comparing said first data structure to a plurality of a second data structure to obtain a comparison result, the plurality of said second data structure representing clusters of web page access patterns of a set of users having a known user profile attribute in common;

evaluating based on said comparison result the plurality of said second data structure and said first data structure to identify a second data structure matching the web page access pattern of the first data structure; and

assigning said unknown user profile attribute of said test user from the matching second data structure to said test user;

wherein the known user profile attribute in common of the sample data [[sets]] set of users corresponds to the unknown user profile attribute of said test user.

29. (Canceled)

30. (Currently Amended): A processor readable storage medium, comprising:

processor readable program code embodied on said processor readable storage medium, said processor readable program code for programming a processor to perform a method for extrapolating user profile information from user web page access patterns, the method comprising:

detecting a set of web pages accessed by a test user having an unknown user profile attribute;

mapping at least a subset of said detected web pages to a first data structure, said first data structure representing a web page access pattern of said test user;

comparing said first data structure to a plurality of a second data structure to obtain a comparison result, the plurality of said second data structure representing

Application No. 10/033,586

clusters of web page access patterns of a sample data set of users having a known user profile attribute in common;

evaluating based on said comparison result the plurality of said second data structure and said first data structure to identify a second data structure matching the web page access pattern of the first data structure; and

assigning said unknown user profile attribute of said test user from the matching second data structure to said test user;

wherein the known user profile attribute in common of the sample data [[sets]] set of users corresponds to the unknown user profile attribute of said test user.

Claims 31-34. (Canceled)

35. (New): The method of claim 5, wherein said demographic information is age information of the test user.

36. (New): The method of claim 2, wherein the multi-dimensional vectors are weighted to dampen disparity between how many time web pages are access by different users.

37. (New): The apparatus of claim 28, wherein said first and second data structures are multi-dimensional vectors, wherein each dimension of said first and said second multi-dimensional vectors corresponds to a different web page, and wherein each dimension of said second data structure corresponds to an average of user multi-dimensional vectors in its corresponding cluster.

38. (New): The apparatus of claim 37, wherein said comparing further comprises determining a distance between said multi-dimensional vectors.

Application No. 10/033,586

39. (New): The apparatus of claim 38, wherein said determining further comprises computing a cosine of an angle between said multi-dimensional vectors.

40. (New): The apparatus of claim 37, wherein said unknown user profile attribute is demographic information.

41. (New): The apparatus of claim 40, wherein said demographic information is gender information of the test user.

42. (New): The apparatus of claim 40, wherein said demographic information is age information of the test user.

43. (New): The apparatus of claim 37, wherein the multi-dimensional vectors are weighted to dampen disparity between how many time web pages are access by different users.

44. (New): The processor readable storage medium of claim 30, wherein said first and second data structures are multi-dimensional vectors, wherein each dimension of said first and said second multi-dimensional vectors corresponds to a different web page, and wherein each dimension of said second data structure corresponds to an average of user multi-dimensional vectors in its corresponding cluster.

45. (New): The processor readable storage medium of claim 44, wherein said comparing further comprises determining a distance between said multi-dimensional vectors.

Application No. 10/033,586

46. (New): The processor readable storage medium of claim 45, wherein said determining further comprises computing a cosine of an angle between said multi-dimensional vectors.

47. (New): The processor readable storage medium of claim 44, wherein said unknown user profile attribute is demographic information.

48. (New): The processor readable storage medium of claim 47, wherein said demographic information is gender information of the test user.

49. (New): The processor readable storage medium of claim 47, wherein said demographic information is age information of the test user.

50. (New): The processor readable storage medium of claim 44, wherein the multi-dimensional vectors are weighted to dampen disparity between how many time web pages are access by different users.